

Modale partikels en vertaling

Willy Vandeweghe (voorheen Hogeschool Gent)

Dutch, like German, is characterized by its prolific use of modal particles, a category which is much less present in languages like English and French. This paper reports on the results of an investigation into the differences between original Dutch and Dutch translated from English or French, for which data from two parallel corpora were used. The hypothesis, that translated Dutch has a lower density of modal particles, due to a lack of stimulus from the source text, was not unequivocally borne out.

1. Inleiding: partikelrijke en partikelarme talen

Vanuit taaltypologie en contrastieve linguïstiek (CL) is al langer bekend dat er talen zijn die kwistig gebruikmaken van adverbiale (en andere) partikels voor het toevoegen van pragmatische betekenselementen aan de zin, en andere die dat veel minder doen. Het Nederlands en het Duits gelden binnen de West-Europese talen als ‘partikelrijke’ talen. Het ligt dan ook in de lijn der verwachtingen dat er op dit front verlies optreedt wanneer uit het Nederlands vertaald wordt naar partikelarmere talen als het Frans of het Engels. Voorts is het aannemelijk dat vertaald Nederlands met deze talen als brontaal, een geringere partikelbezetting vertoont dan origineel Nederlands, zeker wanneer het gaat over modale partikels waarvan sowieso al aangenomen kan worden dat hun betekenisbijdrage meer in het pragmatische vlak ligt.¹

In deze bijdrage worden enkele eerste onderzoeksresultaten voor de modale partikels (MP's) *wel* en *maar* gepresenteerd om deze hypothese te bevestigen of te ontcrachten. De mogelijkheden voor het onderzoek naar dit soort verschillen in taalgebruik zijn behoorlijk toegenomen doordat er steeds meer meertalige digitale corpora ter beschikking komen die ook Nederlands tekstmateriaal bevatten. Voor we de geobserveerde fenomenen presenteren en op zoek gaan naar verklaringen, is het nuttig iets te zeggen over meertalige corpora.

2. Nuttige corpora bij de omgang met meertaligheid

2.1 Soorten corpora voor vertaalwetenschap

Net als de contrastief linguïst kan de vertaalwetenschapper zijn profijt doen met digitale corpora waarin een vergelijkingscomponent zit ingebouwd. Er zijn twee soorten corpora die de vertaalwetenschap met name interesseren: *parallele* en *vergelijkbare* corpora (cf. Baker 1993, 1995, 1996). De *parallele* of *vertaalcorpora* bevatten teksten met hun vertaling. Om die bruikbaar te maken voor praktische doeleinden – en ook voor onderzoek – is het belangrijk de vertaaleenheden aan elkaar te koppelen, wat gebeurt met behulp van *alignering*. Daarbij worden corresponderende zinnen, of vertaaleenheden op een lager niveau, aan elkaar gelinkt, waardoor ze bij doorzoeking van het corpus samen gepresenteerd worden. Figuur 1 is een voorbeeld van alignering in een tekst van een toeristische website (Vandeweghe 2005, 169): het voorbeeld laat zien dat corresponderende zinnen uit origineel en vertaling aan elkaar gekoppeld zijn. Dat gaat soms zin per zin, maar als de volgorde verschilt, kan de koppeling chiaistisch zijn (cf. F4-F5 // N5-N4). Ook nulcorrespondentie is mogelijk, waarbij een zin in de ene taal geen corresponderende zin heeft in de andere, maar die treffen we niet aan in het voorbeeld, evenmin als een-tot-veel- of veel-tot-eencorrespondentie.

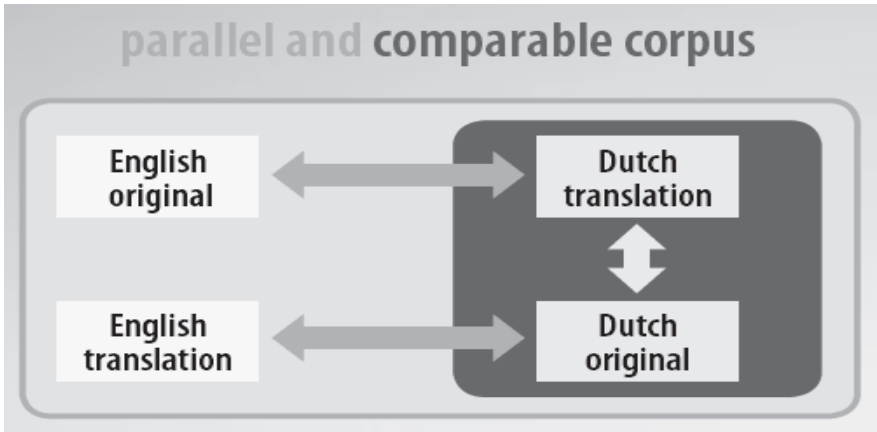
A voir à Mons	F1 — N1	Bezienswaardigheden in Bergen
<i>L'hôtel de ville</i>	F2 — N2	<i>Stadhuis</i>
Le bâtiment	F3 — N3	Het stadhuis
construit en 1458,	F4 — N4	heeft een gotische gevel.
présente une façade gothique.	F5 — N5	en werd gebouwd in 1458
En 1715, un campanile y est adjoint,	F6 — N6	In 1715 werd er een klokkentorentje aan toegevoegd
sans nuire à l'unité de l'édifice.	F7 — N7	maar dit kwam de eenheid van het gebouw niet ten goede.
Le temps qui passe y est mesuré de manière originale.	F8 — N8	Wel wordt de tijd er op een originele manier voorgesteld.

Figuur 1 – Alignering

Vergelijkbare corpora bevatten teksten in eenzelfde taal, zowel origineel geschreven als vertaalde teksten, die uit vergelijkbare tekstsoorten en genres afkomstig zijn, of over vergelijkbare thema's gaan, maar die verder niet corresponderen op het micro-niveau.

Vaak zijn multilinguale corpora allebei: een parallel corpus dat gebalanceerd is, dat wil zeggen waarin de geselecteerde teksten evenredig verdeeld zijn over tekstsoorten en thema's, en dat beide vertaalrichtingen omvat, is tegelijk ook een vergelijkbaar corpus. Zo bevat een parallel corpus Engels-Nederlands/Nederlands-Engels dat opgebouwd is uit romans en vertalingen ervan, zowel origineel NL (in

de vertaalrichting → EN) als vertaald NL (in de vertaalrichting → NL). Die Nederlandse teksten, vertaalde en originele, vormen dan een vergelijkbaar corpus, dankzij het feit dat ze hetzelfde genre bestrijken.



Figuur 2 – Parallel en vergelijkbaar corpus

2.2 Meertalige corpora met Nederlands als spiltal: corpus-Niemegeers en DPC

In het kader van haar contrastief-vertaalkundige studie van modale partikels heeft Sofie Niemegeers (Gent) een parallel én (dus) vergelijkbaar corpus NE-EN/EN-NL aangelegd, bestaande uit krantenartikels, fictioneel proza en theaterteksten. Hieronder worden enkele resultaten gepresenteerd op grond van een pilootstudie over een gedeelte van dit corpus. Dit corpus is enkel voor onderzoeksdoeleinden toegankelijk, in tegenstelling tot het door de TST-centrale van de Nederlandse Taalunie vrij te geven DPC, wat staat voor ‘Dutch Parallel Corpus’. Het gaat hierbij om een parallel corpus Nederlands-Engels-Frans van tien miljoen woorden, tot stand gekomen binnen het STEVIN-programma dat wordt gefinancierd door de Nederlandse en Vlaamse overheden (<http://www.stevin-tst.org>). Dit programma was erop gericht een impuls te geven aan de totstandkoming van technologische taal- en vertaalvoorzieningen voor het Nederlands. Het parallelle corpus, geannoteerd en op zinsniveau gealigneerd, heeft Nederlands als spiltal.² Het is samengesteld aan de KULeuven (campus Kortrijk) en het departement Vertaalkunde van de Hogeschool Gent, en wordt gepresenteerd in Paulussen e.a. (2006). Een website met meer informatie is te vinden op <http://www.kuleuven-kortrijk.be/dpc>.

3. Eigenschappen van vertaalde tekst

Traditioneel houdt de vertaalwetenschap zich bezig met de relatie tussen bron- en doelttekst, om van daaruit inzicht te krijgen in de mechanismen die het vertaalproces – dat een keuzeproces is – beïnvloeden. Een typische uiting daarvan is de belangstelling voor ‘verschuivingen’: aanpassingen in de vertaling waarbij afgeveken wordt van een woord-voor-woord-weergave (bijvoorbeeld *Ik ben hem niet vergeten* > *I remember him*; *les habits que contenait ma valise* > *de kleren in mijn koffer*, cf. Vandeweghe 2005, 91-110). In de loop van de jaren ’70-’80 is het perspectief echter verschoven naar de ‘doel’-pool van de vertaalrelatie, en is er steeds meer belangstelling voor vertaling (het product) binnen de context waar ze deel van uitmaakt, namelijk die van de doelcultuur (cf. Toury 1995). Het is vanuit die verschuiving te verklaren dat de schijnwerper nu ook gericht werd op de vergelijking tussen vertaalde taal en originele taal.

Een eigenschap van vertaalde taal is dat die typische kenmerken vertoont als gevolg van interferentie met de brontaal. De ‘intertaal’ (*interlanguage*) die hier het gevolg van is, wordt meestal negatief beoordeeld, wat tot uiting komt in de term *translationese*. Het gaat daarbij om afwijkingen van het normale taalgebruik zoals die worden aangetroffen in vertaalde teksten, en die te wijten zijn aan invloed van de brontaal. Het kan gaan om fouten tegen *lexicon* of *grammatica* van de doeltaal (cf. NL *eventueel* / FR *éventuellement* vs. EN *eventually*; FR *librairie* ‘boekhandel’ vs. EN *library* ‘bibliotheek’; NL *raar* ‘vreemd’ vs. FR of EN *rare* ‘zeldzaam’; overdadig gebruik van passieve constructies in NL vertaling uit het Engels).

Die intertaal kan echter ook vanuit een meer neutraal standpunt bekeken worden, als een gegeven dat **inherent** is aan vertaling. Er is nog niet veel onderzoek naar gedaan, maar een mogelijke onderzoeksvraag in dat verband is: verschilt vertaalde taal van niet-vertaalde taal op het punt van keuze van altemnerende constructies? Voor het Engels is een dergelijke vergelijking gemaakt voor het gebruik van het verbindingswoord *that* in bijzinnen: zoals bekend kent het Engels een alternantie tussen (1a) en (1b):

- (1a) He said that he was tired.
- (1b) He said \emptyset he was tired.

In hun onderzoek zijn Olohan & Baker (2000) nagegaan of die alternantie ook te maken kan hebben met de factor vertaling. Er blijkt inderdaad een significant verschil te zijn tussen vertaald en origineel Engels, in die zin dat de constructie met *that* verhoudingsgewijs heel wat frequenter is in vertaald Engels. Hoe dit te verklaren is, is een andere vraag, maar er dienen zich een paar verklaringssporen aan:

1. Het is een algemeen gegeven dat vertalers gemakkelijk hun toevlucht nemen tot **explicitering** van zaken die in de brontekst impliciet blijven, gewoon omdat

ze voor hun publiek zo helder mogelijk willen zijn. Een typisch voorbeeld hiervan is de vertaling van culturele referenties, ook soms aangeduid als *realia*. Aangezien de doeltaallezer niet altijd geacht kan worden vertrouwd te zijn met de cultuur waarin de brontekst tot stand gekomen is, kan de vertaler oordelen dat explicitering noodzakelijk is. Voorbeeld: FR *Le Monde* wordt in de NL vertaling: *de Franse kwaliteitskrant Le Monde*. De constructie met *that* is eveneens explicieter dan die zonder, zodat de keuze voor deze constructie gelinkt kan worden aan de neiging van vertalers om explicieter te zijn dan gebruikelijk is in gewone, niet-vertaalde teksten.

2. In dit specifieke geval is de omgekeerde verklaring waarschijnlijker, nl. dat de vertaler zich voor de Engelse constructie laat leiden door die in de brontaal: de meeste talen (Spaans, Frans, Nederlands, ...) hebben niet de alternantie die het Engels heeft, dus in de aangeboden zin is een expliciet voegwoord aanwezig dat de vertaler gewoon meeneemt in zijn Engels, een geval dus van **interferentie** met de brontaal (in de niet-evaluatieve betekenis van de term, cf. *supra*).
3. Een factor die bij vertaling ook speelt, is de tendens tot **normalisering**: de tendens om zich te conformeren aan patronen en praktijken die typisch zijn voor de doeltaal. Deze wat conservatieve, conformerende, houding ter vermindering van risico kan de vertaler ertoe brengen voor de ‘meest verbreide’ en dus ‘veiligste’ constructie te kiezen. Het is een manier om *translationese* te vermijden, wat in sommige gevallen tot hypercorrectie kan leiden. Normalisering in vertaling leidt tot terughoudendheid met creatieve nieuwvormingen. Zo bleek uit een studie van woorden op -ish dat die in vertaald Engels minder voorkwamen dan in origineel Engels, waar termen als *actorish* or *caterpillarish* wel kunnen. Omgekeerd echter waren er meer kleurnamen op -ish in het vertaalde Engels, en dat dan ten koste van de meer idiomatische kleurensynoniemen van het type *hazel* (Olohan 2004, 108-122). Minder negatief geconnoteerde termen voor ‘normalisering’ zijn ‘idiomatisering’ en ‘naturalisering’. Daarmee doelt men meer op de aanpassing die voortvloeit uit de tendens bij vertalers om zich te conformeren aan wat zij beschouwen als de typische patronen, idiomaten en normen die gelden in de doeltaal, volgens ‘the law of growing standardization’ (Toury 1995, 267-274).

Het onderzoek naar *that* / \emptyset illustreert dat vergelijkbare corpora de onderzoeker in staat stellen originele en vertaalde taal als variëteiten met elkaar te vergelijken. Vertalingen worden binnen deze context wel eens gekwalificeerd als een *third code*, een tussenvorm (*hybride*) waarvan de taalkenmerken zich onderscheiden van zowel bron- als doeltaal (Frawley 1984).

Het onderzoek van deze *third code* wil dus te weten komen wat de kenmerken zijn van de taalvariant die eigen is aan vertalingen, hoe die kenmerken gespreid zijn, en met welke relatieve frequentie ze voorkomen. Deze benadering heeft ge-

leid tot een onderzoeksprogramma waarbij men zich toespitst op *vertaaltuniversalia*. Hiermee zijn universele kenmerken bedoeld van vertaalde teksten (in welke doel-taal ook), kenmerken dus die gelden onafhankelijk van de brontaal die aan de vertaling ten grondslag ligt. Voorbeelden: vereenvoudigingen, inkorten van zinnen, opheffen van dubbelzinnigheid, toevoegen van expliciterende informatie, vermijden van herhalingen, simplificatie. De vergelijking van taalvariëteiten biedt indirect inzicht in vertaalbeslissingen door de vertaler(s).

4. Vertaalkundige corpusstudie naar Nederlandse modale partikels

4.1 Pilootstudie van Niemegeers (2007)

Voor het Nederlands zijn nog maar weinig resultaten voorhanden van onderzoek waarin origineel en vertaald NL vergeleken wordt. Niemegeers (2007) doet verslag van een pilootstudie ondernomen in het kader van haar doctoraatsproject waarin ze nagaat hoe de modale partikels zich gedragen in een vertaalcontext: zowel bij vertaling van NL in het EN – partikelarm, net als het FR – als in de vergelijking tussen vertaald (VN) en origineel Nederlands (ON). Eerst een voorbeeld:

- (2a) NL *Luister maar.* (*maar* [ter verzachting van?] imperatief)
- (2b) EN *You better listen.*
- (3a) Wie denken jullie *wel* dat jullie zijn? (*wel* [als markering van een] retorische vraag)
- (3b) Who do you think you are?

In voorbeeld (2) correspondeert met het NL partikel geen EN partikel, maar de hele zin wordt weergegeven door een constructie met het bijwoord *better*. In voorbeeld (3) is er nul-correspondentie: geen spoor van NL *wel* in het Engels. Toch krijgt het partikel soms een vertaalcorrespondentie met partikel, zoals EN *just* voor NL *maar* in voorbeeld (4):

- (4) Vertel *maar.* (NL bronzin uit het corpus)
'Just tell me.' (EN vertaling uit het corpus)

Het corpus van Niemegeers – voor deze pilootstudie samengesteld op basis van fictioneel proza en theaterteksten – is tegelijk parallel (en bidirectioneel: EN→NL en NL→EN) en vergelijkbaar: de NL bronteksten voor de EN vertaling vormen een vergelijkingscorpus voor de teksten in vertaald NL. Tabel 1 geeft de cijfers van het voor de pilootstudie gebruikte deelcorpus:

Vertaalrichting	NL-EN	EN-NL
Totaal aantal woorden	703.675	617.216
Subcorpus 'theaterstukken'	253.946	299.309
Subcorpus 'romans'	449.729	317.907

Tabel 1: aantal woorden uit het corpus ingezet voor de pilootstudie

De vergelijking tussen VN en ON leverde op dat het totale aantal voorkomens van de MP's *maar* en *wel* opmerkelijk hoger ligt in VN dan in ON. De cijfers worden gegeven in tabel 2:

	VN	Aantal per 10.000 wdn	ON	Aantal per 10.000 wdn
Totaal aantal woorden	308.737 (47 %)		347.850 (53 %)	
MP <i>maar</i>	546	18	402	12
MP <i>wel</i>	718	23	468	13

Tabel 2: aantal voorkomens van maar en wel in VN en ON

Het mogelijke interferentie-effect bij elementen die typisch zijn voor een bepaalde taal – zoals de modale partikels in het NL – is door Tirkkonen-Condit (2004, 177-8) geformuleerd als de *Unique Items Hypothesis*. 'Unieke items' zijn lexicale of grammaticale elementen die specifiek zijn voor een taal, en ontbreken in andere talen. Die hypothese voorspelt dat door het gebrek aan beschikbare tegenhangers en dus stimuli in de brontaal dergelijke elementen minder vaak in de vertaling zullen opduiken:

Every language has linguistic elements that are unique in the sense that they lack straightforward linguistic counterparts in other languages. [...] Since they are not similarly manifested in the source language, it is to be expected that they do not readily suggest themselves as translation equivalents, as there is no obvious linguistic stimulus for them in the source text [...] In other words, linguistic elements that are 'unique' in this sense would have lower frequencies in translated texts than in originally produced texts (Tirkkonen-Condit 2004, 177-178).

Het markante nu is dat we net het omgekeerde vaststellen bij de MP's in Nederlands dat uit het EN vertaald is. Een mogelijke verklaring moet wellicht gezocht worden in de **normaliseringstendens** bij vertaling: juist omdat de vertaler weet dat partikels bijdragen tot de idiomatiek van het NL, kan hij/zij geneigd zijn wat gulter met partikels te strooien dan het geval zou zijn in originele taal, in een reactie

van distantiëring ten opzichte van de partikelarme brontaal. Voor de omgekeerde vertaalrichting geven de cijfers van Niemegeers (2007) wel een effect dat als het correlaat van het Unique Items-principe gezien kan worden: hoe partikelarm ook, toch blijkt vertaald Engels (VE) significant meer gebruik te maken van de zeldzaam voorhanden MP-equivalenten, in dit geval het EN just als vertaling voor NL maar. Dat dit significant meer voorkomt in VE lijkt nog maar eens te wijzen op **brontaalinterferentie**: de aanwezigheid van een stimulus in de bron bevordert het gebruik van bepaalde items in de doeltaal.

Bij deze resultaten moeten we de volgende relaterende opmerking maken: het gaat om een al met al beperkt corpus van 1,3 miljoen woorden, waarvan het Nederlandse deel (vertaald en origineel) een goede 600.000 woorden beslaat (656.587), met bovendien een beperkte representativiteit qua taalparen (enkel EN-NL) en genre: louter fictionele teksten. Om na te gaan of deze cijfers en verhoudingen bevestigd worden door ander materiaal, is gebruikgemaakt van voorlopig vrijgegeven DPC-materiaal, met Nederlandse, Engelse en Franse data.

4.2 Resultaten uit prerelease DPC

4.2.1 Materiaal en werkwijze

Het corpus uit de DPC prereleases is qua omvang sterk vergelijkbaar met het corpus-Niemegeers: een 1,3 miljoen woorden, waarvan het Nederlandse deel (vertaald en origineel) uit een goede 600.000 woorden bestaat (636.347).

Vertaalrichting	NL-EN	EN-NL	NL-FR	FR-NL	TOTAAL
Totaal aantal woorden					1.288.131
Nederlands	145.677	206.002	95.106	191.897	638.682
Vreemde taal	157.237	198.950	102.590	193.007	651.784

Tabel 3: aantal woorden uit het DPC-corpus

Het DPC heeft echter twee voordelen ten opzichte van het corpus-Niemegeers: behalve Engels bevat het ook Frans materiaal, en naast ‘literatuur’ bevat het ook andere tekstsoorten: administratieve teksten, instructieve teksten, teksten voor externe communicatie en journalistieke teksten.

Met ‘literatuur’ worden zowel fictie (romans, korte verhalen) als non-fictie (essays, (auto)biografieën, populair-wetenschappelijke teksten) bedoeld. ‘Journalistieke’ teksten zijn interessant omdat het om levendige en – net als bij ‘literatuur’ – creatieve taal gaat. Een beperking hier is dat het in hoofdzaak om Belgisch materiaal gaat, te weten de kwaliteitskranten *De Morgen* en *De Standaard*, en de weekbladen *Knack* en *Trends*. ‘Instructief’ zijn vooral teksten uit handboeken voor apparaten en software, procedurebeschrijvingen en dergelijke meer. ‘Administratief’

gaat vrij breed, van strikt administratieve teksten als jaarboeken, vergaderingsverslagen en zakelijke correspondentie, tot wetgeving, parlementaire debatten en speeches van politici, die ook tot het zakelijke genre behoren. ‘Externe communicatie’ ten slotte betreft documenten waarin organisaties, projecten en evenementen zich presenteren aan de buitenwereld, algemeen-informerende documenten, promotiemateriaal, persteksten enz.

Voor dit kleine onderzoekje zijn dezelfde partikels onder de loep genomen als in de pilootstudie van Sofie Niemegeers. Behalve naar Nederlands vertaald uit het Engels is ook naar vertaling uit het Frans gekeken. Het meest toegankelijk zijn de cijfergegevens over *wel*, omdat dit altijd een adverbiale status heeft (een uitzondering als *wel* en *wee* daargelaten), en zo goed als altijd een modale lading. Een zoekopdracht voor *maar* levert een rijke oogst op, maar die bevat in hoofdzaak attestaties van het nevenschikkende voegwoord, wat voor het hier behandelde thema niet ter zake doet. Omdat de woordsoort-tags in het door mij gebruikte zoekprogramma ParaConc (Barlow 2002, 2008) in een *query* niet gespecificeerd konden worden, heb ik hier met de hand een uitfiltering gemaakt om enkel de adverbiale attestaties over te houden. In de bespreking beperk ik me tot de ruwe cijfergegevens, ook al omdat er uiteindelijk niet zoveel relevante attestaties overbleven. In de cijfers die volgen, is ‘literatuur’ apart genomen, omdat dat het best te vergelijken is met dat van het corpus-Niemegeers.

4.2.2 Resultaten

Wordt het resultaat van de pilootstudie nu bevestigd door de DPC-data? Hier moeten we nuanceren. Niemegeers constateerde dat de modale partikels frequenter voorkwamen in vertaald Nederlands. In ons materiaal echter, samengevat in tabel 4, geldt het omgekeerde, weliswaar met de uitzondering van Nederlands met Engels als brontaal binnen de rubriek ‘literatuur’. Daar sluit de hoge ranking op de lijst van frequente woorden (61e plaats) perfect aan bij de tendens in het materiaal-Niemegeers. Bij adverbiaal partikel (modaal en scalair) *maar* (enkel voor de vier niet-literaire tekstsoorten) is er tussen brontaal- en vertaald Nederlands nauwelijks verschil in frequentie: 52 hits voor EN>NL (op 80.071 woorden is dat 0,064 % of 6 voorkomens per 10.000 woorden), 63 hits voor NL>EN (op 93.824 woorden, is dat 0,067 % of 6,7 per 10.000). Omdat zoals gezegd de adverbiale gebruikswijzen van *maar* via ParaConc niet systematisch te isoleren zijn, hebben we de tel- en rangschikkingsfunctie die dit programma bevat, enkel kunnen toepassen op het *wel*-materiaal. Die gegevens geven we weer in tabel 4.

		aantal woorden			WEL		Rang	
		Absol			Absol	Per		
					10.000			
		NL	NL	VT				
N-E	Adm	93.824	14.994 >	15.521	15	1,1	70	73
	ExtC		36.353 >	39.965	18	0,5		
	Instr		13.060 >	14.195	11	0,8		
	Jour		29.417 >	31.137	75	2,5		
	Liter	51.853	51.853 >	/56.419	48	0,9	97	
	TOT	145.677	145.677 >	157.237	167	1,1		
#	drs. Antoinet Brink, secretaris IVN		+11.560					
	%	7,9						
E-N	Adm	80.071	22.281 <	21.575	4	0,1	130	69
	ExtC		14.164 <	14.143	7	0,5		
	Instr		10.045 <	9.832	1	0,1		
	Jour		33.581 <	32.688	38	1,1		
	Liter	125.931	125.931 <	120.712	238	1,8	61	
	TOT	206.002	206.002 <	198.950	288	1,4		
#	Abs	+7.052						
	%	3,5						
N-F	Adm	77.678	18.052 >	18.962	14	0,8	73	72
	ExtC		23.466 >	27.119	14	0,6		
	Instr		5.501 >	5.625	1	0,5		
	Jour		30.659 >	32.191	68	2,2		
	Liter	17.428	17.428 >	18.693	23	1,3	80	
	TOT	95.106	95.106 >	102.590	120	1,3		
#	Abs	+7.484						
	%	7,8						
F-N	Adm	61.672	20.956 <	21.124	6	0,2	198	132
	ExtC		25.494 <	26.707	12	0,4		
	Instr		283 <	317	-	0		
	Jour		14.939 <	15.098	10	0,6		
	Liter	130.225	130.225 <	129.761	110	8,5	135	
	TOT	191.897	191.897 <	193.007	138	0,7		
#	Abs	-1.110						
	%	0,5						

Tabel 4: DPC cijfergegevens over wel en totaal aantal woorden

Toelichting: de kolom in cursief geeft de cijfers voor het NL weer (als bron- of doeltaal), de haken ‘<’ en ‘>’ duiden op de vertaalrichting. Met ‘rang’ wordt verwezen naar de plaats in de frequentierangschikking voor woorden. ‘+’ en ‘-’ wijzen op resp. ‘toename’ en ‘vermindering’.

Wat eventuele verschillen tussen de tekstsoorten betreft: meestal scoren Journal en Liter het hoogst wat het aantal *wel*-tokens per 10.000 woorden betreft, met Journal FR>NL als opvallende uitzondering, evenals (in mindere mate) Liter NL>EN: slechts 0,9 *wel*-tokens per 10.000 woorden. Markant is dat voor Admin het brontaal-Nederlands systematisch hoger scoort op densiteit van *wel*-voorkomens dan het vertaalde Nederlands: rond de 1 op 10.000 voor het brontaal-Nederlands, max. 0,2 op 10.000 voor de vertaalde variant. Het meest stabiel is de rangschikking voor Ext Comm: telkens in de buurt van 0,5 *wel*-voorkomens per 10.000, zowel in origineel als in vertaald Nederlands.

Er is overigens nog iets uitzonderlijks aan het FR>NL vertaald materiaal: we noteren hier een netto vermindering van het aantal woorden in de vertaalslag. Dat is iets wat ingaat tegen de algemene tendens bij vertalingen. In het standaardgeval is de vertaling dan ook langer dan het origineel, wat de vertaalrichting ook moge zijn, wellicht om de simpele reden dat vertalers als geboren ‘uitleggers’ het expliciteren in het bloed hebben.³ De uitzonderlijke cijfers voor het FR>NL materiaal kunnen overigens heel goed berusten op toeval: ze gelden alleen voor de niet-literatuur, bij literatuur is er juist een verhoging van het aantal woorden. Dat laatste strookt met mijn ervaring met literaire vertaalscripties waar ik kon observeren dat FR>NL vertaling zo goed als altijd met een toename van het aantal woorden gepaard gaat. Als verder onderzoek op ander materiaal echter een tendens aan het licht zou brengen, dan moeten de hypothesen over de universaliteit van het expliciteringsbeginsel bij vertaling danig bijgesteld worden.

5. Besluit

Wanneer we originele en vertaalde taal als twee variëteiten onderzoeken, iets wat dankzij de corpusmethode op steeds representatievere basis kan gebeuren, dan blijken er interessante verschillen uit de vergelijking. Er kan zeker niet zomaar gezegd worden dat vertaald Nederlands, bij gebrek aan stimulus vanuit de brontaal, een lagere densiteit van modale partikels vertoont, zoals onze uitgangshypothese aangaf. Wanneer vanaf 2010 het integrale DPC-corpus beschikbaar is, en wanneer ook andere corpora ter beschikking komen, zullen deze en andere hypothesen natuurlijk op veel ruimer materiaal getoetst kunnen worden, wat niet alleen de statistische waarde van de cijfergegevens zal verhogen, maar ook de zoektocht naar verklaringen haalbaarder zal maken.

Noten

1. Voor focuspartikels als ook, *enkel*, *pas* heeft het Engels meestal wel een equivalent van een vergelijkbare categorie (*also*, *too*, *only*), bij modale partikels (*maar*, *eens*, *even*) is dat zelden het geval.

2. Een deel van de data verwerkt in dit artikel werd geselecteerd uit het Dutch Parallel Corpus.
3. 'Explicitering' is een van de vertaaluniversalia (dit zijn kenmerken die eigen zijn aan vertaling, ongeacht de talen die erbij betrokken zijn) waarover in de vertaalwetenschappelijke literatuur het minst controverserig bestaat.

Bibliografie

- Baker, Mona, 'Corpus linguistics and translation studies: Implications and applications.' M. Baker et al. (red.), *Text and Technology. In honour of John Sinclair*. Amsterdam and Philadelphia, 1993, 233-250
- Baker, Mona, 'Corpora in Translation Studies: An Overview and Some Suggestions for Future Research'. *Target*, 7, 2 (1995), 223-43.
- Baker, Mona, 'Corpus-based translation studies: The challenges that lie ahead'. H. Somers (ed.) *Terminology, LSP, and Translation: Studies in language engineering, in honour of Juan C. Sager*. Amsterdam, 1996, 175-86.
- Barlow, G. Michael, 'ParaConc: Concordance Software for Multilingual Parallel Corpora'. *Proceedings of the Third International Conference on Language Resources and Evaluation. Workshop on Language Resources in Translation Work and Research*. Las Palmas, Spain, 2002, 20-24.
- Barlow, G. Michael, *ParaConc and Parallel Corpora in Contrastive and Translation Studies*. Houston, 2008, 110.
- Frawley, William, 'Prolegomenon to a Theory of Translation.' W. Frawley (red.), *Translation: Literary, Linguistic, and Philosophical Perspectives*. London & Toronto, 1984, 159-175
- Niemegeers, Sofie, 'The Dutch modal particles *maar* and *wel* and their English correspondences: evidence from a parallel and comparable corpus.' *Lezing EST congress, Ljubljana* 2007.
- Olohan, Maeve, *Introducing Corpora in Translation Studies*. London/New York, 2004.
- Olohan, Maeve & Mona Baker, 'Reporting that in translated English: Evidence for subconscious processes of explicitation?' *Across Languages and cultures* 1/2(2000), 141-158.
- Tirkkonen-Condit, S., 'Unique items: over- or under- represented in translated language?' *Translation Universals: Do they exist?*, A. Mauranen and P. Kuusimäki. Amsterdam/ Philadelphia, 2004, 177-186.
- Toury, Gideon, *Descriptive translation studies and beyond*. Amsterdam/Philadelphia, 1995.
- Paulussen, Hans et al., 'Dutch Parallel Corpus: a multifunctional and multilingual corpus'. *Cahiers de l'Institut de Linguistique de Louvain, CILL, Louvain-La-Neuve*, 32.1-4 (2006), 269-285
- Vandeweghe, Willy, *Duoteksten. Inleiding tot vertaling en vertaalstudie*. Gent, 2005 (gereviseerde 3e druk, 2008).